

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: HUSE, WILLIAM D.

(ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF  
HETEROMERIC RECEPTORS

(iii) NUMBER OF SEQUENCES: 75

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
(B) STREET: 444 SO. FLOWER STREET, SUITE 200  
(C) CITY: LOS ANGELES  
(D) STATE: CALIFORNIA  
(E) COUNTRY: UNITED STATES  
(F) ZIP: 90071

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: CAMPBELL, CATHRYN A.  
(B) REGISTRATION NUMBER: 31,815  
(C) REFERENCE/DOCKET NUMBER: P31 8882

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 619-535-9001  
(B) TELEFAX: 619-535-8949

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7445 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60  
 ATAGCTAAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120  
 CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTTCCAGACA CGGTACTTTA 180  
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240  
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300  
 TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAACGCG ATATTGAAG 360  
 TCTTTCGGGC TTCTCTTAA TCTTTTIGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	480
TTTGAGGGGG	ATTCAATGAA	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	540
AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT	600
GGTTTTTATC	GTCGTCTGGT	AAACGAGGGT	TATGATAGTG	TTGCTCTTAC	TATGCCTCGT	660
AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAATGTG	GTATTCCTAA	ATCTCAACTG	720
ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GTCCTGACTG	GTATAATGAG	CCAGTCTTA	AAATCGCATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG	960
AATATCCGGT	TCTTGTCAAG	ATTACTTTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCCGGCT	AAGTAACATG	GAGCAGGTCC	CGGATTTCGA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCCTTCC	CGTCTTTCGT	TTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCGT	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TAACTCCCT	GCAAGCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTGGGCG	ATGGTTGTTG	TCATTGTCCG	CGCAACTATC	GGTATCAAGC	TGTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGTCCTTTT	GGAGCCTTTT	1560
TTTTTGAGA	TTTTCAACGT	GAAAAAATTA	TTATTCGCAA	TTCCTTTAGT	TGTTCCTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCGGGGCT	ATACTTATAT	CAACCCCTCTC	GACGGCACTT	ATCCGGCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAACTG	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTGAGA	GACTGCGCTT	TCCATTCTGG	GTTTAATGAA	2220
GATCCATTCC	TTGTGAATA	TCAAGGCCAA	TGGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460

GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCCGTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTOAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCCTTTG	GGTATTCGGT	2880
TATTATTGCG	TTTCCTCGGT	TTCCCTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTTCATT	GTTTGTGTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTG TG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTCAGGG	TGTTCAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTG	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTTATTT	GTAAGTGCCA	AATTAGGCTC	TGGAAAGACG	3240
CTCGTTAGCG	TTGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCCGAGGT	TCGCTAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGGCGTAC	TTGGTTTAAT	3480
ACCGGTTCTT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTG	TCTTATTACT	GGCTCGAAAA	TCCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCTTAT	TTATCACACG	GTCCGTTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTC	ACGCGTTCTT	3960
TGTCCTGCCA	TTGGATTTCG	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTIAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCAGTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTATAG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCCTC	TGCGCGATTT	4320
TGTAACITGG	TATTCAAAGC	AATCAGGCCA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTTCAT	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTTATTTG	4440
TGTTTTACGT	GCTAATAAAT	TTGATATGGT	TGTTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500

TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TACTCAAAC	4620
TTTTAAATT	AATAACGTT	GGGCAAAGGA	TTTAATACGA	GTTGTCGAAT	TGTTTGTAAC	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTCTACTG	TTGATTGGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTGGGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCCGCGAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC	AGAAAGGTTT	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAATAAAT	CCATTTGAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACCGT	TAATTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATT	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	5580
TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	5700
ATTGGGGTGA	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC	TCTTGTTCCT	AACTGGAACA	AACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTGGCC	GATTTGGGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGCGGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCCT	6000
GGGGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATT	GGCAGCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCGTC	ACTTGGCACT	GGCGTCGTTT	TTACAACGTC	6240
GTGACTGGGA	AAACCCTGGC	GTTACCCAAG	CTTTGTACAT	GGAGAAAATA	AAGTGAAACA	6300
AAGCACTATT	GCACTGGCAC	TCTTACCGTT	ACCGTTACTG	TTTACCCCTG	TGACAAAAGC	6360
CGCCAGGTC	CAGCTGCTCG	AGTCAGGCCT	ATTGTGCCCA	GGGGATTGTA	CTAGTGATC	6420
CTAGGCTGAA	GGCGATGACC	CTGCTAAGGC	TGCATTCAAT	AGTTTACAGG	CAAGTGCTAC	6480
TGAGTACATT	GGCTACGCTT	GGGCTATGGT	AGTAGTTATA	GTTGGTGCTA	CCATAGGGAT	6540

TAAATTATTC AAAAAGTTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGCCCGCACC	6600
GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCTTTGC CTGCTTTCCG	6660
GCACCAGAAG CCGTGCCGGA AAGCTGGCTG GAGTGCGATC TTCCTGAGGC CGATACGGTC	6720
GTCGTCCCCT CAAACTGGCA GATGCACGGT TACGATGCGC CCATCTACAC CAACGTAACC	6780
TATCCCATT ACGTCAATCC GCGTTTGTG CCCACGGAGA ATCCGACGGG TTGTTACTCG	6840
CTCACATTTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTGTAT	6900
GGCGTTCCTA TTGGTTAAAA AATGAGCTGA TTAAACAAAA ATTTAACGCG AATTTTAACA	6960
AAATATTAAC GTTTACAATT TAAATATTTG CTTATACAAT CTTCTGTGTT TTGGGGCTTT	7020
TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTTAGGATTA CCGTTCATCG	7080
ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTTGTA GATCTCTCAA	7140
AAATAGCTAC CCTCTCCGGC ATTAATTTAT CAGCTAGAAC GGTGAATAT CATATTGATG	7200
GTGATTGAC TGTCTCCGGC CTTTCTCAGC CTTTGAATC TTTACCTACA CATTACTCAG	7260
GCATTGCATT TAAAATATAT GAGGGTCTA AAAATTTTAA TCCTTGCGTT GAAATAAAGG	7320
CTTCTCCCGC AAAAGTATTA CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT	7380
GCTCTGAGGC TTTATTGCTT AATTTTGCTA ATGTTTGCC TTGCCTGTAT GATTTATTGG	7440
ACGTT	7445

## (2) INFORMATION FOR SEQ ID NO:2:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7317 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTGCGGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTGCGAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAACATGT TGAGCTACAG CACCAGATTG AGCAATTAAAG CTCTAAGCCA	240
TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGTCT GGTTCGCTTT GAAGCTCGAA TTAACGCGG ATATTTGAAG	360
TCTTTGCGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGGTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCTCGT	660
AATTCCTTTT GCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720

ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTC	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGAAC	1080
GTCTGCGCCT CGTTCGGGCT AAGTAACATG GAGCAGGTGC CGGATTTCCA CACAAATTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTATAGT TATTCTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAAATG AAACCTTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTG CTACCCTGCT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCGGTGA	1440
TGCGTGGGCG ATGGTGTGTT TCAATTGTCG CGCAACTATC GGTATCAAGC TGTTTAAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGAGAG TTTTCAACGT GAAAAAATA TTATTGCGAA TTCCTTTAGT TGTTCTTTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAAATCA	1680
TTTACTAACG TCTGGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACCTCAGT TTACGGTACA	1800
TGGGTTCTTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAAC TGTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GCGGCTCTG AGGCTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGCTGGCGG CTCTGAGGGA GCGGTTCCG GTGGTGGCTC TGGTTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACCTGATT CTGTCCGTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGGCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760

TTCCGTGGTG	TCTTTGCCGT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCCGTA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCCGT	2880
TATTATTGCG	TTTCCTCGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTCATT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTGTT	GGTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTCAGGG	TGTTCACTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTG	3120
TCTCTGTAAA	GGCTGCTATT	TTCAATTTTG	ACGTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGGAAAGACG	3240
CTCGTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAG	CGGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTG	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCTTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTG	ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTTGC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTIATTCA	CTCACATATA	TGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATGTTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATTT	4320
TGTAACITGG	TATTCAAAGC	AATCAGGCCA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTGAT	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTTATTTG	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAAAAT	AATAACGTTG	GGGCAAAGGA	TTTAATACGA	GTTGTGGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTCTACTG	TTGATTTGCC	4800

AACTGACCAG ATATTGATTG AGGGTTTGGT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGGT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCCGT ATTTTAAATG GCGATGTTTT	4980
AGGGTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGAAATAAT CCATTTCAGA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGGTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTAGC CGCCCTGTAG CGGCGCATT AAGCGGCGCG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCGGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTCTTCCA AACTGGAACA AACTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCG GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGACGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATT A GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC	6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCGTGAT	6300
GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAATT CTAGAAGCGG TCACTTGCCA	6360
CTGGCCGTG TTTTACAACG TCGTGAAGTG GAAAACCGTG GCGTTACCCA AGCTTAATCG	6420
CCTTGAGAAA TTCCCTTTTC CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	6480
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTTT GCCTGGTTTC CGGCACCAGA	6540
AGCGGTGCCG GAAAGCTGGC TGGAGTGCGA TCTTCCTGAG GCCGATAGGG TCGTCGTCCC	6600
CTCAAAGTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT	6660
TACGGTCAAT CCGCCGTTTG TTCCCACCGA GAATCCGACG GGTGTTTACT CGCTCACATT	6720
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTTG ATGGCGTTCC	6780
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA	6840



ACGTTTACAA	TTTAAATATT	TGCTTATACA	ATCTTCCTGT	TTTTGGGGCT	TTTCTGATTA	6900
TCAACCGGGG	TACATATGAT	TGACATGCTA	GTTTTACGAT	TACCGTTCAT	CGATTCTCTT	6960
GTTTGCTCCA	GACTCTCAGG	CAATGACCTG	ATAGCCTTTG	TAGATCTCTC	AAAAATAGCT	7020
ACCCTCTCCG	GCATTAATTT	ATCAGCTAGA	ACGGTTGAAT	ATCATATTGA	TGGTGATTTG	7080
ACTGTCTCCG	GCCTTTCTCA	CCCTTTTGAA	TCTTTACCTA	CACATTACTC	AGGCATTGCA	7140
TTTAAATAT	ATGAGGGTTC	TAAAAATTTT	TATCCTTGCG	TTGAAATAAA	GGCTTCTCCC	7200
GCAAAAGTAT	TACAGGGTCA	TAATGTTTTT	GGTACAACCG	ATTTAGCTTT	ATGCTCTGAG	7260
GCTTTATTGC	TTAATTTTGC	TAATTCTTTG	CCTTGCCTGT	ATGATTTATT	GGATGTT	7317

## (2) INFORMATION FOR SEQ ID NO:3:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7729 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA	CTATTAGTAG	AATTGATGCC	ACCTTTTCAG	CTCGCGCCCC	AAATGAAAAT	60
ATAGCTAAAC	AGGTTATTGA	CCATTTGCGA	AATGTATCTA	ATGGTCAAAC	TAAATCTACT	120
CGTTCGAGA	ATTGGGAATC	AACTGTTACA	TGGAATGAAA	CTTCCAGACA	CCGTACTTTA	180
GTTGCATATT	TAAAACATGT	TGAGCTACAG	CACCAGATTC	AGCAATTAAG	CTCTAAGCCA	240
TCTGCAAAAA	TGACCTCTTA	TCAAAAGGAG	CAATTAAAGG	TACTCTCTAA	TCCTGACCTG	300
TTGGAGTTTG	CTTCCGGTCT	GGTTCGCTTT	GAAGCTCGAA	TTAAAACGCG	ATATTGGAAG	360
TCTTTCCGGC	TTCTCTTAA	TCTTTTTGAT	GCAATCCGCT	TTGCTTCTGA	CTATAATAGT	420
CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	480
TTTGAGGGGG	ATTCAATGAA	TATTTATGAC	GATTCCGAG	TATTGGACGC	TATCCAGTCT	540
AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT	600
GGTTTTTATC	GTCGTCTGGT	AAACGAGGGT	TATGATAGTG	TTGCTCTTAC	TATGCCTCGT	660
AATTCCTTTT	GGCGTATGT	ATCTGCATTA	GTTGAATGTG	GTATTCCTAA	ATCTCAACTG	720
ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CGGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GTCCTGACTG	GTATAATGAG	CCAGTTCCTA	AAATCCGATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACCTTGAT	TTGGGTAATG	960
AATATCCGGT	TCTTGCAAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCGGGCT	AAGTAACATG	GAGCAGGTCC	CGGATTTCCA	CACAAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT	1200

CAAAGATGAG	TGTTTTAGTG	TATTCTTTCG	CCTCTTTCGT	TTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACGCTCGT	TCCGATGCTG	TCTTTCGGTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTGA	1440
TGCGTGGGCG	ATGGTTGTTG	TCATTGTCCG	CGCAACTATC	GGTATCAAGC	TGTTAAGAA	1500
ATTACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGAGA	TTTCAACGT	GAAAAAATTA	TTATTGCGAA	TTCCTTTAGT	TGTTCTTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAAATCA	1680
TTTACTAACC	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AACTCAGTG	TTACGGTACA	1800
TGGGTTCCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCGGGGCT	ATACTTATAT	CAACCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAAC	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTGACA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTGG	TTTGTGAATA	TCAAGGCCAA	TGGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGGG	GCTCTGGTGG	TGGTCTGGT	GCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGGCG	TACAGTCTGA	CGCTAAAGGC	AACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTGG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCGGTCA	ATATTACCT	TCCCTCCCTC	AATCGGTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTGGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	ACTTCTTTTG	GGTATTCCGT	2880
TATTATTGGG	TTTCCTCGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTCAAT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATCTTGTG	GTTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTACAGG	TGTTACAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTG	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGAAAGACC	3240

CTCGTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GCCTTCAAAA	CCTCCCGCAA	GTCCGGAGGT	TCGCTAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCC	3600
CGTTCTGCA	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTAGAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AAOGCCTTAT	TTATCAGACG	GTCCGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTC	ACGCGTTCCT	3960
TGTCTTGCGA	TTGGATTTGC	ATCAGGATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATT	4320
TGTAACCTGG	TATTCAAAGC	AATCAGGCGA	ATCGTTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTICAT	CTGACGTTAA	ACCTGAAAAAT	CTACGCAATT	TCTTTATTTT	4440
TGTTTTACGT	GCTAATAAAT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAAAAT	AATAACGTTT	GGGCAAAGGA	TTTAATACGA	GTTGTCGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACGTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTCCGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCCGGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTGAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280

TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GCGGTACCGT TCCTGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATT AAGCGCGGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA AACTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTATAAG GGATTTTGCC GATTTCCGAA CCACCATCAA	5880
ACAGGATTTT CGCCTGCTGG GGCAAACAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT	6000
GGGCGCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATT AAGCAACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCGTC ACTTGGCACT GGCCGTCGTT TTACAACGTC	6240
GTGACTGGGA AAACCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAACA	6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT	6360
CCAGCTGCTC GAGTCGGTCT TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC	6420
AGCGGCCCTG GGCTGCCTGG TCAAGACTAA TTCCCCGAAC CGGTGACGGT GTCGTGGAAC	6480
TCAGGCGCCC TGACCAGCGG CGTGCACACC TTCCCGGCTG TCCTACAGTC CTCAGGACTC	6540
TACTCCCTCA GCAGCGTGGT GACCGTGCCC TCCAGCAGCT TGGGCACCCA GACCTACATC	6600
TGCAACGTGA ATCACAAGCC CAGCAACACC AAGGTGGACA AGAAAGCAGA GCCCAAATCT	6660
TGTACTAGTG GATCCTACCC GTACGACGTT CCGGACTACG CTCTTTAGGC TGAAGGGGAT	6720
GACCCCTGCTA AGGCTGCATT CAATAGTTTA CAGGCAAGTG CTAAGTAGTA CATTGGCTAC	6780
GCTTGGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAAATT ATTCAAAAAG	6840
TTTACGAGCA AGGCTTCTTA AGCAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC	6900
AGTTGCGCAG CCTGAATGGC GAATGGCGCT TTGCCTGGTT TCCGGCACCA GAAGCGGTGC	6960
CGGAAAGCTG GCTGGAGTGC GATCTTCCTG AGGCCGATAC GGTCTGCTGC CCTCAAAC	7020
GGCAGATGCA CGGTTACGAT GCGCCCATCT ACACCAACGT AACCTATCCG ATTACGGTCA	7080
ATCGGCGGTT TGTTCACAGG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTAAATGTTG	7140
ATGAAAGCTG GCTACAGGAA GGCCAGACGC GAATTATTTT TGATGGCGTT CATTATTGGT	7200
AAAAAATGAG CTGATTAAAC AAAAAATTTAA CGCGAATTTT AACAAAATAT TAACGTTTAC	7260
AATTTAAATA TTTGCTTATA CAATCTTCCT GTTTTGGGG CTTTCTGAT TATCAACCGG	7320

GGTACATATG ATTGACATGC TAGTTTTACG ATTACCGTTC ATCGATTCTC TTGTTTGCTC 7380  
 CAGACTCTCA GGCAATGACC TGATAGCCTT TGTAGATCTC TCAAAAATAG CTACCCTCTC 7440  
 CGGCATTAAT TTATCAGCTA GAACGGTTGA ATATCATATT GATGGTGATT TGACTGTCTC 7500  
 CGGCCTTTCT CAGCCTTTTG AATCCTTACC TACACATTAC TCAGGCATTG CATTAAAAAT 7560  
 ATATGAGGGT TCTAAAAAT TTTATCCTTG CGTTGAAATA AAGGCTTCTC CCGCAAAAGT 7620  
 ATTACAGGGT CATAATGTTT TTGGTACAAC CGATTTAGCT TTATGCTCTG AGGCTTTATT 7680  
 GCTTAATTTT GCTAATTCTT TGCCTTGCCT GTATGATTTA TTGGACGTT 7729

(2) INFORMATION FOR SEQ ID NO:4:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7557 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGOC ACCTTTTCAG CTCGGCCCCC AAATGAAAAT 60  
 ATAGCTAAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120  
 CGTTCGCAGA ATTGGGAATC AACTGTTACA TGAATGAAA CTTCAGACA CCGTACTTTA 180  
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240  
 TCCGCAAAA TGACCTCTTA TCAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300  
 TTGGAGTTTG CTTCGGTCT GGTTCGCTTT GAAGGTCGAA TAAAACCGG ATATTTGAAG 360  
 TCTTCGGGC TTCCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420  
 CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTGCT TTTCTGAACT GTTTAAAGCA 480  
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540  
 AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT 600  
 GGTTTTIATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660  
 AATTCCTTTT GCGGTATGT ATCTGCATTA GTTGAATGTG GTATTCTTAA ATCTCAACTG 720  
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780  
 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTCTTAA AAATCGCATA AGGTAATTCA 840  
 CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTAGTCTG TCTGGTGTTT 900  
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTCAT TTGGGTAATG 960  
 AATATCCGGT TCTTGTCAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT CGGCCTGGTC 1020  
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080  
 GTCTGGCCCT CGTTCGGCT AAGTAACATG GAGCAGGTCT CCGATTTCGA CACAATTTAT 1140  
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTCGGGC TTGGTATAAT CGCTGGGGGT 1200  
 CAAAGATGAG TGTTTTAGTG TATCTTTCTG CCTCTTCTG TTTAGGTTGG TGCCTTCGTA 1260

GTGGCATTAC GTATTTTACC CGTTTAAATGG AAACCTTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCCGA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCCG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCGTTTT GGAGCCTTTT	1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCGAA TTCCTTTAGT TGTTCCTTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACG TCTGGAAAGA CQACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACCTCAGTG TTACGGTACA	1800
TGGGTTCCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCGGGGCT ATACTTATAT CAACCCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTCAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GTTCCGAAA TAGGCAAGGG GCATTAACTG TTTATACGGG CACTGTACT	2100
CAAGGCACTG ACCCCGTAA AACTTATAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTCTG TTTGTGAATA TCAAGGCCAA TGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GCGGTTCCG GTGGTGGCTC TGGTCCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGGCG TACAGTCTGA CGCTAAAGGC AAACCTGATT CTGTGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTCTTTTG GGTATTCGGT	2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CCGCTATCTG CTTACTTTTC	2940
TTAAAAGGG CTTCGGTAAG ATAGCTATTG CCTGTTTCTT GCTCTTATTA TTGGGCTTAA	3000
CTCAATTCTT GTGGGTATC TCTCTGATAT TAGCGCTCAA TTACCCTCTG ACTTTGTTCA	3060
GGGTGTTTCA TTAATTCTCC CGTCTAATGC GCTTCCCTGT TTTTATGTTA TTCTCTCTGT	3120
AAAGGCTGCT ATTTTCATTT TTGACGTTAA ACAAAAATC GTTTCTTATT TGGATTGGGA	3180
TAAATAATAT GGCTGTTTAT TTTGTAACCT GCAAATTAGG CTCTGGAAAG ACGCTCGTTA	3240
GCGTTGGTAA GATTCAGGAT AAAATTGTAG CTGGGTGCAA AATAGCAACT AATCTTGATT	3300

TAAGGCTTCA	AAAGCTCCCG	CAAGTCGGGA	GGTTCGCTAA	AAGGCCTCGC	GTTCTTAGAA	3360
TACCGGATAA	GCCTTCTATA	TCTGATTTCG	TTGCTATTGG	GCGCGGTAAT	GATTCCCTACG	3420
ATGAAAATAA	AAACGGGTTG	CTTGTTCTCG	ATGAGTCCGG	TACTTGGTTT	AATACCCGTT	3480
CTTGGAATGA	TAAGCAAACA	CAGCCGATTA	TTGATTGGTT	TCTACATGCT	CGTAAATTAG	3540
GATGGGATAT	TATTTTCTT	GTTCAGGACT	TATCTATTGT	TGATAAACAG	GCGCGTTCG	3600
CATTAGCTGA	ACATGTTGTT	TATTGTCGTC	GTCTGGACAG	AATTACTTTA	CCTTTTGTG	3660
GTACTTTATA	TTCTCTTATT	ACTGGCTCGA	AAATGCCTCT	GCCTAAATTA	CATGTTGGCG	3720
TTGTAAATA	TGGCGATTCT	CAATTAAGCC	CTACTGTTGA	GCGTTGGCTT	TATACTGGTA	3780
AGAATTTGTA	TAACGCATAT	GATACTAAAC	AGGCTTTTTC	TAGTAATTAT	GATTCCGGTG	3840
TTTATTCTTA	TTTAACGCCT	TATTTATCAC	ACGGTCGGTA	TTTCAAACCA	TTAAATTTAG	3900
GTCAGAAGAT	GAAGCTTACT	AAAATATATT	TGAAAAAGTT	TTCACGCGTT	CTTTGTCTTG	3960
CGATTGGATT	TGCATCAGCA	TTTACATATA	TTTATATAAC	CCAACCTAAG	CCGGAGGTTA	4020
AAAAGGTAGT	CTCTCAGACC	TATGATTTTG	ATAAATTCAC	TATTGACTCT	TCTCAGCGTC	4080
TTAATCTAAG	CTATCGCTAT	GTTTTCAAGG	ATTCTAAGGG	AAAATTAATT	AATAGCGACG	4140
ATTTACAGAA	GCAAGGTIAT	TCACTCACAT	ATATTGATTT	ATGTAAGTTT	TCCATTAAAA	4200
AAGGTAATTC	AAATGAAATT	GTTAAATGTA	ATTAAATTTG	TTTTCTTGAT	GTTTGTTC	4260
TCATCTTCTT	TTGCTCAGGT	AATTGAAATG	AATAATTCGC	CTCTGCGCGA	TTTTGTAACT	4320
TGGTATTCAA	AGCAATCAGG	CGAATCCGTT	ATTGTTTCTC	CCGATGTAAA	AGGTACTGTT	4380
ACTGTATATT	CATCTGACGT	TAAACCTGAA	AATCTACGCA	ATTTCTTTAT	TTCTGTTTTA	4440
CGTGCTAATA	ATTTTGATAT	GGTTGGTTCA	ATTCCTTCCA	TAATTCAGAA	GTATAATCCA	4500
AACAATCAGG	ATTATATTGA	TGAATTGCCA	TCATCTGATA	ATCAGGAATA	TGATGATAAT	4560
TCCGCTCCTT	CTGGTGGTTT	CTTTGTTCCG	CAAAATGATA	ATGTTACTCA	AACTTTTAAA	4620
ATTAATAACG	TTCGGGCAAA	GGATTTAATA	CGAGTTGTCG	AATTGTTTGT	AAAGTCTAAT	4680
ACTTCTAAAT	CCTCAAATGT	ATTATCTATT	GACGGCTCTA	ATCTATTAGT	TGTTAGTGCA	4740
CCTAAAGATA	TTTTAGATAA	CCTTCCTCAA	TTCTTTTCTA	CTGTTGATTT	GCCAACTGAC	4800
CAGATATTGA	TTGAGGGTTT	GATATTTGAG	GTTCAGCAAG	GTGATGCTTT	AGATTTTTC	4860
TTTGCTGCTG	GCTCTCAGCG	TGGCACTGTT	GCAGGCGGTG	TTAATACTGA	CCGCCTCACC	4920
TCTGTTTTAT	CTTCTGCTGG	TGGTTCGTTT	GGTATTTTAA	ATGGCGATGT	TTTAGGGCTA	4980
TCAGTTCGCG	CATTAAAGAC	TAATAGCCAT	TCAAAAATAT	TGTCTGTGCC	ACGTATTCTT	5040
ACGCTTTCAG	GTCAGAAGGG	TTCTATCTCT	GTTGGCCAGA	ATGTCCCTTT	TATTACTGGT	5100
CGTGTGACTG	GTGAATCTGC	CAATGTAAAT	AATCCATTTT	AGACGATTGA	GCGTCAAAAT	5160
GTAGGTATTT	CCATGAGCGT	TTTCCTGTT	GCAATGGCTG	GCGGTAATAT	TGTTCTGGAT	5220
ATTACCAGCA	AGGCCGATAG	TTTGAGTTCT	TCTACTCAGG	CAAGTGATGT	TATTACTAAT	5280
CAAAGAAGTA	TTGCTACAAC	GGTTAATTTG	CGTGATGGAC	AGACTCTTTT	ACTCGGTGGC	5340

CTCACTGATT ATAAAAACAC TTCTCAAGAT TCTGGCGTAC CGTTCCTGTC TAAAATCCCT	5400
TTAATCGGGC TCCTGTTTAG CTCCGGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG	5460
CTCGTCAAAG CAACCATAGT ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGTGTGGT	5520
GGTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT	5580
CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT	5640
CCCTTAGGG TTCCGATTTA GTGCTTIACG GCACCTCGAC CCCAAAAAC TTGATTGEG	5700
TGATGGTTCA CGTAGTGGG CATCGCCCTG ATAGACGGTT TTTCCGCCCT TGACGTTGGA	5760
GTCCACGTTT TTTAATAGTG GACTCTTGTT CCAAACCTGGA ACAACACTCA ACCCTATCTC	5820
GGGCTATTCT TTTGATTAT AAGGATTTT GCCGATTTCC GAACGACCAT CAAACAGGAT	5880
TTTCGCCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG	5940
GTGAAGGGCA ATCAGCTGTT GCGCGTCTCG CTGGTGAAAA GAAAAACCAC CCTGGCGCCC	6000
AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG	6060
GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA	6120
TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG	6180
CGGATAACAA TTTCACACGC CAAGGAGACA GTCAATAATGA AATACCTATT GCCTACGGCA	6240
GCCGCTGGAT TGTATTACT CGCTGCCCAA CCAGCCATGG CCGAGCTCTT CCCGCCATCT	6300
GATGAGCAGT TGAAATCTGG AAGTGCCTCT GTTGTGTGCC TGCTGAATAA CTTCTATCCC	6360
AGAGAGGCCA AAGTACAGTG GAAGGTGGAT AACGCCCTCC AATCGGGTAA CTCACAGGAG	6420
AGTGTACAG AGCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGCTG	6480
AGCAAAGCAG ACTACGAGAA ACACAAAGTC TACGCCTGCG AAGTCACCCA TCAGGGCCTG	6540
AGCTCGCCCG TCACAAAGAG CTTCAACAGG GGAGAGTGT CTAGAACCGG TCACTTGCCA	6600
CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA AGCTTAATCG	6660
CCTTGCAGAA TTCCCTTTCC CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	6720
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTTT GCTGGTTTC CGGCACCAGA	6780
AGCGGTGCCG GAAAGCTGGC TGGAGTGGCA TCTTCCTGAG GCCGATACGG TCGTCGTCCC	6840
CTCAAACTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT	6900
TACGGTCAAT CCGCCGTTTG TTCCACCGGA GAATCCGACG GGTGTTACT CGCTCACATT	6960
TAATGTTGAT GAAAGCTGGC TACAGGAAGC CCAGACCGCA ATTATTTTG ATGGCGTTCC	7020
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATT	7080
ACGTTTACAA TTIAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGGT TTTCTGATT	7140
TCAACCGGGG TAGATATGAT TGACATGCTA GTTTTACGAT TACCGTTTAT CGATTCTCTT	7200
GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT	7260
ACCTCTCCG GCATTAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TCGTGAATTC	7320
ACTGTCTCCG GCCTTTCTCA CCCTTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA	7380



TTTAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC	7440
GCAAAAGTAT TACAGGGTCA TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG	7500
GCTTTATTGC TTAATTTTGC TAATCTTTG CCTTGCCTGT ATGATTTATT GGATGTT	7557

## (2) INFORMATION FOR SEQ ID NO:5:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8118 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTGSCA AATCTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TCGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CAACAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGGTCT GGTTCCGCTT GAAGCTCGAA TTAACACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTIGAT GCAATCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TCGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGCTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTIT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCGGGT AAGTAACATG GAGCAGGTG CGGATTCGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATCTTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AACTTTCCTC ATGAAAAAGT CTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCGCA AAAGCGGCCT TAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TCCGTGGGCG ATGTTGTTG TCATTGTGG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500

ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT 1560  
 TTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCGAA TTCCTTTAGT TGTTCCTTTC 1620  
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA 1680  
 TTTACTAACG TCTGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740  
 CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACCTCAGTG TTACGGTACA 1800  
 TGGGTTCTTA TTGGGCTTGC TATCOCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860  
 TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920  
 ATTCCGGGCT ATAATTATAT CAACCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA 1980  
 AACCCCGCTA ATCTAATCC TTCTCTTGA GAGTCTCAGC CTCTTAATAC TTTCATGTTT 2040  
 CAGAATAATA GGTTCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT 2100  
 CAAGGCACTG ACCCCGTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160  
 TATGACGCTT ACTGGAACGG TAAATTCAGA CAOTGGGCTT TCCATTCTGG CTTTAATGAA 2220  
 GATCCATTCTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT 2280  
 GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GCGGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340  
 GCGGTTCTG AGGGTGGCGG CTCTGAGGGA GCGGTTCTG CTGGTGGCTC TGGTTCCGGT 2400  
 GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460  
 GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACCTGATT CTGTCGCTAC TGATTACGGT 2520  
 GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580  
 GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640  
 TTAATGAATA ATTCGGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700  
 TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760  
 TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820  
 TTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCCTTTG GGTATTCGGT 2880  
 TATTATTGCG TTTCCTCGGT TTCTTCTGG TAACCTTGTG CGGCTATCTG GTTACTTTTC 2940  
 TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCAAT GTTCTTGCT CTTATTATTG 3000  
 GGCTTAACTC AATTCCTGTG GGTATCTCT CTGATATTAG CGCTCAATTA CCTCTGACT 3060  
 TTGTTACGGG TGTTCACTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC 3120  
 TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAATCGTT TCTTATTGCG 3180  
 ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACGGCA AATTAGGCTC TGGAAAGACG 3240  
 CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300  
 CTTGATTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAC GCCTCGCGTT 3360  
 CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT 3420  
 TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTCCGGTAC TTGGTTTAAAT 3480  
 ACCCGTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT 3540

AAATTAGGAT	GGGATATTAT	TTTTCTTGT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCGTCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCCTTAT	TTATCACACG	GTCCGGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTC	ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTTGC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTTAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTTCT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATTT	4320
TGTAACCTTG	TATCAAAGC	AATCAGGCGA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTCAT	CTGACGTTAA	AECTGAAAAT	CTACGCAATT	TCTTTATTTT	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGTTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAATTT	AATAACGTTT	GGGCAAAGGA	TTTAATACGA	GTTGTCGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AAGTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACGTGTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTGTTCCGGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCCGCGAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGGTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CGGCTGTGAT	TCCAACGACG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGCGGCATTA	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	5580

TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TGCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTAGGGTTT	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	5700
ATTTGGGTGA	TGTTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATACTGGAC	TCTTGTTCCA	AACTGGAACA	AACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTTGCC	GATTTCGGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCCT	6000
GGCGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATTA	GGCACCCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCCAA	CGAGACAGTC	ATAATGAAAT	ACCTATTGCC	6240
TACGGCAGCC	GCTGGATTGT	TATTACTCGG	TCCCAACCA	GCCATGGCCG	AGCTCTTCCC	6300
GCCATCTGAT	GAGCAGTTGA	AATCTGGAAC	TGCTCTGTGT	GTGTGCCTGC	TGAATAACTT	6360
CTATCCCAGA	GAGGCCAAAG	TACAGTGGAA	GGTGGATAAC	GCCCTCCAAT	CGGGTAACTC	6420
CCAGGAGAGT	GTCACAGAGC	AGGACAGCAA	GGACAGCACC	TACAGCCTCA	GCAGCACCCT	6480
GACGCTGAGC	AAAGCAGACT	ACGAGAAACA	CAAAGTCTAC	GCCTGCGAAG	TCACCCATCA	6540
GGGCTGAGC	TGCCCGTCA	CAAAGAGCTT	CAACAGCGGA	GAGTGTCTTA	GAACGCGTCA	6600
CTTGCCACTG	GCCGTCGTTT	TACAACGTCC	TGACTGGGAA	AACCCTGGCG	TTACCCAAGC	6660
TTTGTACATG	GAGAAAATAA	AGTGAAACAA	AGCACTATTG	CACTGGCACT	CTTACCGTTA	6720
CTGTTTACCC	CTGTGGCAAA	AGCCGCCTCC	ACCAAGGGCG	CATCGGTCTT	CCCCCTGGCA	6780
CCCTCTCCA	AGAGCACCTC	TGGGGGCACA	GCGGCCCTGG	GCTGCCTGGT	CAAGACTAAT	6840
TCCCCGAACC	GGTGACGGTG	TCGTGGAAC	CAGGGCCCTT	GACCAGCGGC	GTGCACACCT	6900
TCCCGGCTGT	CCTACAGTCC	TCAGGACTCT	ACTCCCTCAG	CAGCGTGGTG	ACCGTGCCTT	6960
CCAGCAGCTT	GGGCACCCAG	ACCTACATCT	GCAACGTGAA	TCACAAGCCC	AGCAAGACCA	7020
AGGTGGACAA	GAAAGCAGAG	CCCAAATCTT	GTAAGTGTGG	ATCGTACCCG	TACGACGTTT	7080
CGGACTACGC	TTCTTAGGCT	GAAGGCGATG	ACCCTGCTAA	GGCTGCATTG	AATAGTTTAC	7140
AGGCAAGTGC	TACTGAGTAC	ATTGGCTACG	CTTGGGCTAT	GGTAGTAGTT	ATAGTTGGTG	7200
CTACCATAGG	GATTAAATTA	TTCAAAAAGT	TTACGAGCAA	GGCTTCTTAA	GCAATAGCGA	7260
AGAGGCCCCG	ACCGATCGCC	CTTCCCAACA	GTTGCGCAGC	CTGAATCGCG	AATGGCGCTT	7320
TGCCTGGTTT	CCGGCACCAG	AAGCGGTGCC	GGAAAGCTGG	CTGGAGTGGG	ATCTTCCTGA	7380
GGCCGATACG	GTCGTCGTCC	CCTCAAACTG	GCAGATGCAC	GGTTACGATG	CGCCCATCTA	7440
CACCAACGTA	ACCTATCCCA	TTACGGTCAA	TCCGCCGTTT	GTTCCACGGG	AGAATCCGAC	7500
GGGTTGTTAC	TCGCTCACAT	TTAATGTTGA	TGAAAGCTGG	CTACAGGAAG	GCCAGACGGG	7560
AATTATTTTT	GATGGCGTTC	CTATTGGTTA	AAAAATGAGC	TGATTTAACA	AAAATTTAAC	7620

GCGAATTTTA ACAAATATT AACGTTTACA ATTTAAATAT TTGCTTATAC AATCTTCCTG 7680  
 TTTTGGGGC TTTTCTGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTTACGA 7740  
 TTACCGTTCA TCGATTCTCT TGTTTGCTCC AGACTCTCAG GCAATGACCT GATAGCCTTT 7800  
 GTAGATCTCT CAAAAATAGC TACCCTCTCC GGCATTAAAT TATCAGCTAG AACGGTTGAA 7860  
 TATCATATTG ATGGTGATT GACTGTCTCC GGCCTTTCTC ACCCTTTTGA ATCTTTACCT 7920  
 ACACATTACT CAGGCATTGC ATTTAAAATA TATGAGGGTT CTAAAAATTT TTATCCTTGC 7980  
 GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTTT TGGTACAACC 8040  
 GATTAGCTT TATGCTCTGA GGCTTTATTG CTTAATTTTG CTAATTCTTT GCCTTGCCTG 8100  
 TATGATTAT TGGACGTT 8118

(2) INFORMATION FOR SEQ ID NO:6:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc difference
- (B) LOCATION: replace(5, "")
- (D) OTHER INFORMATION: /note= "S REPRESENTS EQUAL MIXTURE OF G AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc difference
- (B) LOCATION: replace(6, "")
- (D) OTHER INFORMATION: /note= "M REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note= "R REPRESENTS EQUAL MIXTURE OF A AND G"

(ix) FEATURE:

- (A) NAME/KEY: misc difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note= "K REPRESENTS EQUAL MIXTURE OF G AND T"

(ix) FEATURE:

- (A) NAME/KEY: misc difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCGAGTCW GG

## (2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCCAGCT GCTCGAGTCT GG

22

## (2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCCAGCT GCTCGAGTCA GG

22

## (2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCCAGCT TCTCGAGTCT GG

22

## (2) INFORMATION FOR SEQ ID NO:10:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCCAGCT TCTCGAGTCA GG

22

## (2) INFORMATION FOR SEQ ID NO:11:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
AGGTCCAACT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:12:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
AGGTCCAACT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:13:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:  
AGGTCCAACT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:14:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:14:  
AGGTCCAACT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:15:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ix) FEATURE:  
(A) NAME/KEY: misc difference  
(B) LOCATION: replace(5..6, "")  
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:  
(A) NAME/KEY: misc difference  
(B) LOCATION: replace(8, "")  
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:  
 (A) NAME/KEY: misc\_difference  
 (B) LOCATION: replace(11, "")  
 (D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:  
 (A) NAME/KEY: misc\_difference  
 (B) LOCATION: replace(20, "")  
 (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:16:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 38 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAAC TA GTAACGGTAA CAGTGGTGCC TTGCCCCA

38

(2) INFORMATION FOR SEQ ID NO:17:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCTTACTA GTACAATCCC TGGGCACAAT

30

(2) INFORMATION FOR SEQ ID NO:18:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCGGA GCTCGTTGTG ACTCAGGAAT CT

32

(2) INFORMATION FOR SEQ ID NO:19:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear



- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
CCAGTTCGGA GCTCGTGTG ACCGAGCCGC CC 32
- (2) INFORMATION FOR SEQ ID NO:20:
- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
CCAGTTCGGA GCTCGTGCTC ACCCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:21:
- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
CCAGTTCGGA GCTCCAGATG ACCCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:22:
- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:22:  
CCAGATGTGA GCTCGTGATG ACCCAGACTC CA 32
- (2) INFORMATION FOR SEQ ID NO:23:
- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:23:  
CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA 32
- (2) INFORMATION FOR SEQ ID NO:24:
- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCAGTTCGGA GCTCGTGATG ACACAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:25:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC

32

(2) INFORMATION FOR SEQ ID NO:26:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGCCGTCTA GAATTACAC TCATTCTGT TGAA

34

(2) INFORMATION FOR SEQ ID NO:27:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC

37

(2) INFORMATION FOR SEQ ID NO:28:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTCAATAGT TTACAGGCAA GTGCTACTGA GTACA

35

## (2) INFORMATION FOR SEQ ID NO:29:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTGGCTACGC TTGGGCTATG GTAGTAGTTA TAGTT

35

## (2) INFORMATION FOR SEQ ID NO:30:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

## (2) INFORMATION FOR SEQ ID NO:31:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACGAGCAAG GCTTCTTA

18

## (2) INFORMATION FOR SEQ ID NO:32:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAACTTTTT GAATAATTT

39

## (2) INFORMATION FOR SEQ ID NO:33:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 36 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:33:  
AATCCCTATG GTAGCAOCAA CTATAACTAC TACCAT

36

(2) INFORMATION FOR SEQ ID NO:34:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:34:  
AGCCCAAGCG TAGCCAATGT ACTCAGTAGC ACTTG

35

(2) INFORMATION FOR SEQ ID NO:35:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:35:  
CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTG

34

(2) INFORMATION FOR SEQ ID NO:36:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:36:  
ATCGCCTTCA GCCTAG

16

(2) INFORMATION FOR SEQ ID NO:37:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:37:  
CATTTTGGCA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:38:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:  
TAGCATTAAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:39:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:  
ATATATTTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:40:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:  
GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:41:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:  
GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT

35

(2) INFORMATION FOR SEQ ID NO:42:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 43 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:  
AAACGACGGC CAGTGCCAAG TGACGGGTGT GAAATTGTTA TCC

43

## (2) INFORMATION FOR SEQ ID NO:43:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 43 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GGCGATTAAG CTTGGGTAAC GCC

43

## (2) INFORMATION FOR SEQ ID NO:44:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG

36

## (2) INFORMATION FOR SEQ ID NO:45:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

## (2) INFORMATION FOR SEQ ID NO:46:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTTAC CCCTGTGACA AAAGCCGCCC AGGTCCAGCT GC

42

## (2) INFORMATION FOR SEQ ID NO:47:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 44 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:47:  
TCGAGTCAGG CCTATTGTGC CCAGGGATTG TACTAGTGGA TCCG

44

(2) INFORMATION FOR SEQ ID NO:48:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:48:  
TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCCTG

38

(2) INFORMATION FOR SEQ ID NO:49:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:49:  
GGCACAATAG GCCTGACTCG AGCAGCTGGA CCAGGGCCGC TT

42

(2) INFORMATION FOR SEQ ID NO:50:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:50:  
TTGTCACAGG GGTAAACAGT AACGGTAACG GTAAGTGTGC CA

42

(2) INFORMATION FOR SEQ ID NO:51:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:51:  
GTGCAATAGT GCTTTGTTTC ACTTTATTTT CTCCATGTAC AA

42

(2) INFORMATION FOR SEQ ID NO:52:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:52:  
TAACGGTAAG AGTGCCAGTG C

21

(2) INFORMATION FOR SEQ ID NO:53:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:53:  
CACCTTCATG AATTCGGCAA GGAGACAGTC AT

32

(2) INFORMATION FOR SEQ ID NO:54:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:54:  
AATTCGCCAA GGAGACAGTC AT

22

(2) INFORMATION FOR SEQ ID NO:55:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:55:  
AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATTGTT

39

(2) INFORMATION FOR SEQ ID NO:56:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:  
ATTACTCGCT GCCCAACCAG CCATGGCCGA GCTCGTGAT

39



## (2) INFORMATION FOR SEQ ID NO:57:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAAT

39

## (2) INFORMATION FOR SEQ ID NO:58:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACGC GTC

13

## (2) INFORMATION FOR SEQ ID NO:59:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 45 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACGGC TTCTAGAATT AACACTCAAT CCTGT

45

## (2) INFORMATION FOR SEQ ID NO:60:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATATCTG GAGTCTGGGT CATCAGGAGC TCGGCCATG

39

## (2) INFORMATION FOR SEQ ID NO:61:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:61:  
GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC

39

(2) INFORMATION FOR SEQ ID NO:62:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:62:  
GTAGGCAATA GGTATTTTCAT TATGACTGTC CTTGGCG

37

(2) INFORMATION FOR SEQ ID NO:63:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:63:  
TGACTGTCTC CTTGGCGTGT GAAATTGTTA

30

(2) INFORMATION FOR SEQ ID NO:64:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 36 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:64:  
TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:65:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:65:  
GCCAGTGCCA AGTGACGCGT TCTA

24

(2) INFORMATION FOR SEQ ID NO:66:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:66:  
ATATATTTTA GTAAGCTTCA TCTTCT 26

(2) INFORMATION FOR SEQ ID NO:67:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:67:  
GACAAAGAAC GCGTGAAAAC TTT 23

(2) INFORMATION FOR SEQ ID NO:68:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 76 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:68:  
CTGAACCTGT CTGGGACCAC AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTAGAGA 60  
CTGGCCTGGC TTCTGC 76

(2) INFORMATION FOR SEQ ID NO:69:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 80 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:69:  
TCGACCGTTG GTAGGAATAA TGCAATTAAT GGAGTAGCTC TAAATTCAGA ATTCATCTAC 60  
ACCCAGTGCA TCCAGTAGCT 80

(2) INFORMATION FOR SEQ ID NO:70:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:70:  
GGTAAACAGT AACGGTAAGA GTGCCAG 27

## (2) INFORMATION FOR SEQ ID NO:71:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 54 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAAGCG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG

54

## (2) INFORMATION FOR SEQ ID NO:72:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 41 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CACCGGTTTCG GGGAATTAGT CTTGACCAGG CAGCCCAGGG C

41

## (2) INFORMATION FOR SEQ ID NO:73:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 51 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTCCACACA TTATACGAGC CGGAAGCATA AAGTGTCAGG CCTGGGGTGC C

51

## (2) INFORMATION FOR SEQ ID NO:74:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGCTCATCA GATGGCGGGA AGAGCTCGGC CATGGCTGGT TG

42

## (2) INFORMATION FOR SEQ ID NO:75:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

78

(x1) SEQUENCE DESCRIPTION; SEQ ID NO:75:  
GAACAGAGTG ACCGAGGGGG CGAGCTGGC CATGGCTGGT TG